





# The Real Preschoolers of Orange County: Early Number Learning in a Diverse Group of Children

Barbara W. Sarnecka<sup>1</sup> , James Negen<sup>2</sup> , Nicole R. Scalise<sup>3</sup> , Meghan C. Goldman<sup>1</sup>, Jeffrey N. Rouder<sup>1</sup> 

[1] Department of Cognitive Sciences, University of California – Irvine, Irvine, CA, USA. [2] Department of Psychology, Durham University, Durham, United Kingdom. [3] Department of Human Development, Washington State University, Pullman, WA, USA.

Journal of Numerical Cognition, 2023, Vol. 9(1), 65–88, <https://doi.org/10.5964/jnc.6577>

Received: 2021-04-16 • Accepted: 2022-10-30 • Published (VoR): 2023-03-31

Handling Editor: Lieven Verschaffel, KU Leuven, Leuven, Belgium

Corresponding Author: Barbara W. Sarnecka, Department of Cognitive Sciences, Social and Behavioral Sciences Gateway, University of California – Irvine, CA 92697-5100, USA. E-mail: [sarnecka@uci.edu](mailto:sarnecka@uci.edu)

Supplementary Materials: Data, Materials [see Index of Supplementary Materials]



## Abstract

The authors assessed a battery of number skills in a sample of over 500 preschoolers, including both monolingual and bilingual/multilingual learners from households at a range of socio-economic levels. Receptive vocabulary was measured in English for all children, and also in Spanish for those who spoke it. The first goal of the study was to describe entailment relations among numeracy skills by analyzing patterns of co-occurrence. Findings indicated that transitive and intransitive counting skills are jointly present when children show understanding of cardinality and that cardinality and knowledge of written number symbols are jointly present when children successfully use number lines. The study's second goal was to describe relations between symbolic numeracy and language context (i.e., monolingual vs. bilingual contexts), separating these from well-documented socio-economic influences such as household income and parental education: Language context had only a modest effect on numeracy, with no differences detectable on most tasks. However, a difference did appear on the scaffolded number-line task, where bilingual learners performed slightly better than monolinguals. The third goal of the study was to find out whether symbolic number knowledge for one subset of children (Spanish/English bilingual learners from low-income households) differed when tested in their home language (Spanish) vs. their language of preschool instruction (English): Findings indicated that children performed as well or better in English than in Spanish for all measures, even when their receptive vocabulary scores in Spanish were higher than in English.

## Keywords

bilingual, children, cardinality, counting, Head Start, number

Children's earliest number skills—skills such as counting out loud, labeling quantities, and recognizing written numerals—undergird later mathematics learning and predict later mathematics performance in both primary and secondary school (Duncan et al., 2007; Watts et al., 2014). However, there is considerable variability in early number skills, and studies of bilingual or multilingual learners are scarce. The goal of the present study was to examine the development of number skills in bilingual<sup>1</sup> or multilingual preschoolers, as well as in monolingual English learners, across a wide socioeconomic range.

When children learn the counting system of their language, they acquire a system for representing exact numbers (Carey, 2009; Sarnecka, 2016). This system is the basis for much of later mathematics (Duncan et al., 2007; Geary et al., 2018; Jordan, Kaplan, Ramineni, & Locuniak, 2009). In early childhood, a few number skills stand out as key. The first



is counting: Children in numerate societies begin learning to count, both intransitively (i.e., reciting the number list) and transitively (i.e., reciting the list while pointing to items one at a time) very early. One study found that toddlers in Australia and Japan begin to recognize the counting list and correct counting procedure sometime between 15 and 18 months of age, and that this recognition was related to the amount of time their parents spent counting with them (Slaughter et al., 2011). Around two years old, many children memorize and recite the beginning of the counting list (e.g., “one, two, three . . .”) just as they recite the beginning of the alphabet (“A, B, C . . .”). At this point, children may know that counting has something to do with quantity (as when one of our toddlers pointed to a tall stack of CDs and said, “onetwothree, onetwothree, onetwothree!” as if to say, “Look at all those CDs!”) but this is not counting in the sense that older children and adults count to determine the exact number of items in a set (Baroody & Price, 1983; Briars & Siegler, 1984; Frye et al., 1989; Fuson, 1988).

At the same time as they are learning to count, children are also learning the cardinal meanings of number words. Children learn the meanings of the first three or four number words one at a time and in order, starting with one and two (Sarnecka & Lee, 2009; Wynn, 1992). This pattern has been shown not only in English, but also in child learners of Spanish, Arabic, Japanese, Mandarin Chinese, Russian, Slovenian, and Tsimane’. (Almoammer et al., 2013; Li et al., 2003; Piantadosi et al., 2014; Sarnecka et al., 2007).

A particularly important step seems to be when children figure out how counting is connected to exact set sizes via the cardinality principle (i.e., that the last word of a count is the exact, cardinal number of the whole set). Children who understand the cardinality principle identify numerosity (as opposed to some other property of sets, such as total summed area) as the dimension of experience that number words pick out (Slusser & Sarnecka, 2011). Similarly, only cardinality-principle-knowers (abbreviated *CP-knowers*) seem to understand that each next word in the list means the addition of one more item to the set (i.e., the successor function, Sarnecka & Carey, 2008; Spaepen et al., 2018), at least for familiar numbers (Cheung et al., 2017; Davidson et al., 2012). However, knowledge of the cardinality principle alone does not guarantee a child’s understanding of the successor function for *all* numbers. It takes CP-knowers approximately two years to learn the specific successor of every number in their count list, and to learn that all numbers have a successor—learning that is aided by their experience with simple  $n+1$  addition facts in early elementary school (Cheung et al., 2017; R. M. Schneider et al., 2020). There is also a strong correlation between understanding the cardinality principle (as tested by the Give-N task) and understanding that the members of two sets of the same number (e.g., five peaches and five apples) can be matched up in one-to-one correspondence with each other (Sarnecka & Wright, 2013). Informally, one could say that when children grasp the cardinality principle, they gain a new conceptual understanding of what numbers are. This understanding is a prerequisite for acquiring the kindergarten and first-grade skills that predict math achievement in later years (Geary et al., 2018; Geary & vanMarle, 2016; Jordan et al., 2006, 2007, 2009; Merkley & Ansari, 2016; Moore et al., 2016).

Distinct from both counting and understanding the meanings of spoken number words, children in literate societies also learn to recognize and name written numerals, and this understanding relates to their later mathematical achievement (Chard et al., 2005; Clarke & Shinn, 2004; Purpura et al., 2013). Written numerals also appear on the number line, which is the most important representation of number-to-space mapping used in early math education (Anobile et al., 2012; Geary et al., 2008; Gunderson et al., 2012; Maertens et al., 2016; Ramani & Siegler, 2008). Number-line estimation tasks are used to probe children’s understanding of the relative magnitudes of exact symbolic numbers, and number-line estimation performance has been related to concurrent and later math achievement (M. Schneider et al., 2018; Siegler, 2016). Moreover, several interventions that focus on improving preschool children’s number-line estimation performance have led to significant improvements on arithmetic and broader math tasks (Cankaya et al., 2014; Honoré & Noël, 2016; Siegler & Ramani, 2009).

Another potential influence on early numeracy is the child’s nonsymbolic, approximate number sense (Libertus & Brannon, 2009; Starr et al., 2013). Long before they learn either spoken or written symbols for numbers, children have access to the innate, nonsymbolic representations of approximate numerical magnitude that humans share with other

---

1) Children this young are more properly called “dual-language learners” or “multiple-language learners” than bilinguals or multilinguals because they are still learning their languages. But because the longer terms are both unwieldy and unfamiliar to many readers, we will use the term ‘bilingual’ here to mean children who are exposed to and learning in more than one language before age 5.

animals (Brannon & Merritt, 2011; Dehaene, 2011; Izard et al., 2009; Xu et al., 2005). These approximate representations allow children to differentiate among sets of varying quantities, such as a pile of three coins and a pile of seven coins. There has been substantial research into the question of how a person's symbolic number skills are related to their approximate number system (ANS) acuity. Meta-analyses have shown that ANS acuity relates to children and adults' performance on symbolic math achievement measures (Chen & Li, 2014; Fazio et al., 2014; M. Schneider et al., 2017). Accordingly, some researchers suggest that children's approximate number representations provide the foundation for their developing symbolic number skills, such that symbolic number understanding is mapped directly onto existing non-symbolic representations (Dehaene, 2001; Feigenson et al., 2004; Piazza et al., 2010). However, other researchers argue that children's understanding of exact symbolic numbers emerges separately from their ANS, and the development of symbolic number knowledge further refines children's approximate representations (Leibovich & Ansari, 2016; Reynvoet & Sasanguie, 2016). Although preschool children's ANS acuity correlates with their understanding of cardinality, written numerals, and symbolic magnitudes in some studies (e.g., Geary & vanMarle, 2016; Mussolin et al., 2012; van Marle et al., 2014), it is possible that this relation may be explained by other related cognitive skills. Indeed, Negen and Sarnecka (2015) demonstrated that many preschoolers do not understand that the verbal instruction to select the group with "more dots" means *a greater number of dots*, rather than for example bigger dots, or dots covering a larger area. Negen and Sarnecka found that after children completed a brief training to clarify the numerical meaning of the term "more," children's ANS acuity was no longer related to their exact number knowledge. Similarly, Wang et al. (2021) demonstrated that experimentally manipulating the ANS acuity of preschoolers who were 4.5 years old and older led to corresponding improvements on their exact math performance, whereas there was no change to the math performance of children younger than 4.5 years old. In sum, the evidence from preschoolers suggests that although the ANS is certainly recruited when children think about numbers (Negen & Sarnecka, 2011), ANS acuity is unlikely to predict symbolic number knowledge in children who do not yet understand how verbal or written numerals represent quantities.

The present study sought to document the performance of a diverse sample of preschool-age children on a range of early numeracy tasks. Although there is a substantial literature on each area (counting, cardinality, nonsymbolic number discrimination, etc.), the variety of tasks used and the small sample sizes in most studies make it difficult to assemble an overall picture of preschool numeracy development or to compare developmental data across studies.

In addition to number tasks, the present study also included measures of children's vocabulary in order to better understand the interplay of language and numeracy in early development. We were particularly interested in whether bilingual learners acquire number concepts on the same timeline as monolingual learners and whether a child's number knowledge in the language of their preschool instruction differs from number knowledge in their home language. On the one hand, one might expect exposure to two languages to be helpful: Switching back and forth between languages might help children learn to control their attention and ignore irrelevant information, both of which are important skills for math learning (Bialystok, 2009, 2011; Bialystok & Martin, 2004; Emmorey et al., 2008; Poulin-Dubois et al., 2011; but see Goldman et al., 2014 for a contrasting result). Similarly, Daubert and Ramani (2019) found that bilingual preschoolers showed higher non-verbal working memory skills than a sample of monolingual preschoolers after controlling for socioeconomic status. Over time, greater working memory skills might allow bilingual learners to better hold and manipulate relevant information in mind while solving math problems (see also Bull & Lee, 2014; Geary & Hoard, 2005).

On the other hand, there are reasons to expect numeracy development to take longer in bilingual learners. A child learning two languages has, in a way, twice as much information to acquire as a child learning only one. Marchand et al. (2020) demonstrated that bilingual children's number word estimations for approximate quantities varied between their languages, suggesting that children store two distinct representations of each number— one in each language. So we might expect bilingual children's math learning to be slightly delayed relative to monolinguals. Several studies have shown that children's number-word and numeral knowledge are closely related to their vocabulary size (LeFevre et al., 2010; Negen & Sarnecka, 2012; Purpura & Napoli, 2015), and bilingual learners typically have smaller vocabularies in each of their languages than monolingual learners of the same age (Bialystok et al., 2010). So we might expect children learning numbers in two languages to take longer than those learning in one. Moreover, if preschool instruction is in English, then children who speak a different home language might in principle miss out on some instruction, so it seems important to consider preschool numeracy development in the context of language development.

Finally, the present study included a large and economically diverse sample of preschoolers. Basic research in cognitive development, as in other areas of psychology, has often relied on convenience samples of participants from middle-class, primarily white and English-monolingual households (Coll, 2015; Henrich et al., 2010; Nielsen et al., 2017). This makes it difficult to translate basic research into educational practices in places like Orange County, where the present studies were conducted. In Orange County, about 1.6 million residents speak only English at home, while almost as many—1.4 million—speak another language at home, with the most common being Spanish (U.S. Census Bureau, 2019; Walters, 2015). As elsewhere in the United States, many Orange County families with young children face economic hardship (Gennetian et al., 2015; U.S. Census Bureau, 2019). Unsurprisingly, this hardship is associated with educational challenges for the children, who come to kindergarten with less knowledge of numbers than their more privileged peers and are four times more likely to show no improvement in these subjects during kindergarten and early first grade (Griffin et al., 1994; Jordan et al., 2006; Jordan & Levine, 2009; National Research Council, 2009; Starkey et al., 2004). The present study was designed to draw a more fine-grained picture of numerical development than has previously been available, particularly for young bilingual learners from economically stressed households. In the present study, we asked children to count (both intransitively reciting the number list and transitively counting objects), to produce sets of a particular number (the Give-N task), to identify the written numerals 1-10, and to place those numerals on a number line. We also tried, with less success, to measure children's nonverbal approximate number system (ANS) acuity, and we measured their receptive vocabulary in English, as well as in Spanish for those children who spoke it.

The study had three goals. The first goal was to explore entailment relations among early numeracy skills by identifying which skills (empirically, not theoretically) seem to form the foundation for other skills, such that they are present by the time the other skills are present. Our measures included symbolic skills (those using spoken and written numerals) as well as a nonsymbolic number comparison (ANS acuity) task where children had to decide which of two dot clouds had “more” dots. The study's second goal was to examine the relation between symbolic number-skill development and language use, comparing bilingual to monolingual learners and separating bilingualism from well-documented socio-economic influences such as household income and parental education. The study's third goal was to find out, for bilingual learners, whether and how children's symbolic number knowledge differed when tested in their home language vs. the language of their preschool instruction.

## Method

### Participants

A total of 566 children were recruited for the study. Of these, 4 had developmental or language delays (as reported by their parent or guardian on the study consent form) and 3 had consent forms where the question about developmental and language delays was left blank. These 7 children were allowed to participate in the study (interact with the researchers and play the games), but their data were excluded from the analysis. Another 13 records were duplicate enrollments where the same child had signed up for the study more than once. These were merged with the original records to create one record per child. An additional 22 children signed up for the study but did not participate, either because they were not at school on any of the testing days or because they did not want to play. Finally, 11 children were either younger than 3 years or older than 6 years at the time of the first testing session. These 11 children were also excluded from the analysis. The resulting sample of 513 children (246 girls, 267 boys) contributed data for the analysis.

These children ranged in age from 3 years, 0 months to 5 years, 11 months ( $M = 4$  years, 5 months;  $SD = 6.8$  months) at the time of the first testing session. At the time of study enrollment, parents were given a demographic questionnaire with questions about home language use, caregiver educational attainment and household income. (The demographic questionnaire and other data collection forms are posted along with the data on the Open Science Framework, see [Supplementary Materials](#)).

### Home Language Use

Parents were asked three questions about the child's home language use: (1) In what language(s) do you and other caregivers speak to your child at home? (2) In what language(s) does your child speak to you and other caregivers at home? (3) Please specify the other language here. Questions 1 and 2 were answered on a 5-point Likert scale with the following values: English only, Mostly English but sometimes another language, Both languages about equally, Mostly another language but some English, Only another language. Based on the parent's responses to these questions, each child participant was assigned a score representing the amount of time that he or she spent interacting in English *outside of school hours*. Scores ranged from 0% English (for children who spoke only another language at home) to 100% English (for children who spoke only English at home). Note that because all of the children attended Head Start or other preschool programs with English instruction, none were monolingual learners of other languages. A score of 0% English represented a child who heard English 0% of the time *outside of school hours*.

### Caregivers' Educational Attainment

The demographic questionnaire also asked parents to "list the caregivers who spend the most time with the child in a typical week (for example: mother, father, grandmother, babysitter, etc.) and mark that person's highest level of education completed." Responses were given on a Likert scale from "1st-11th grade (please write the highest grade completed)" to "Master's or doctoral degree." Each participant was assigned an educational attainment score that reflected the highest level of education attained by any of the child's primary caregivers.

### Household Income

Household income was assessed with the question, "What was the total combined income of all members of the household this past year? (Please include money from jobs, welfare, social security payments, pensions, dividends, etc.\*)" Respondents were free to write in an exact number or choose an income category.

### Ethnicity

The demographic questionnaire asked parents to identify their child's ethnicity (Hispanic or non-Hispanic). Fifty-four percent of children were identified as Hispanic ( $n = 276$ ), 38% were identified as non-Hispanic ( $n = 198$ ), and parents did not report the ethnicity of the remaining 8% ( $n = 39$ ) children. We do not have parent reports of children's race (as distinct from ethnicity), but we believe the racial demographics of the sample to be representative of the communities where the children were recruited. High-SES participants were mainly recruited in the city of Irvine. According to [Data USA \(2022\)](#), the 5 largest racial/ethnic groups in Irvine are Asian, Non-Hispanic (42.8%), White, Non-Hispanic (39.3%), White, Hispanic (8.5%), Two or more groups, Non-Hispanic (4.97%), and Black or African American, Non-Hispanic (2.18%). Low-SES residents came mainly from the cities of Costa Mesa and Santa Ana. The 5 largest racial/ethnic groups in Costa Mesa are White, Non-Hispanic (50.1%), White, Hispanic (21%), Other, Hispanic (12.8%), Asian, Non-Hispanic (8.36%), and Two or more groups, Non-Hispanic (2.86%). The 5 largest racial/ethnic groups in Santa Ana are Other, Hispanic (56.1%), White, Hispanic (20%), Asian, Non-Hispanic (10.9%), White, Non-Hispanic (8.91%), and Two or more groups, Hispanic (1.31%). The high proportion of "Other, Hispanic" likely reflects a widespread view that "Hispanic" is itself a racial identity, distinct from White, Black, and other racial identities.

### Preschools

Research sites included four Head Start programs and seven private pre-kindergarten programs. Head Start is a federally funded program to provide early childhood education to children from low-income households. Accordingly, over 90% of children attending Head Start had annual household incomes below \$30,000, whereas over 90% of children attending private pre-kindergarten programs had annual household incomes of over \$75,000. English was the language of instruction at all of the programs. Although the Head Start teachers were fluent in Spanish, they used Spanish mainly for classroom management and English for curriculum instruction, as their goal was to prepare children for kindergarten programs that would take place entirely in English (see also [Miller, 2016](#)).

## Recruitment

Undergraduate research assistants visited preschools during the morning drop-off and afternoon pick-up hours to invite parents and children to participate in the study. At preschools where many children were Spanish/English bilingual learners, recruitment was done by Spanish/English bilingual research assistants. Consent was obtained in either Spanish or English, as parents preferred. If the parent was unable to fill out the form (e.g., many parents were carrying infants, diaper bags, backpacks, etc.) a research assistant read the questions and recorded the parent's answers. When the person dropping off or picking up the child was not the parent, they were given a consent form to take home for the parent to sign and return later. Children were given a small prize (e.g., a rubber duck or a slinky) at the time of sign-up; no prizes were given during testing.

We followed an inclusive policy in recruiting and testing, meaning that all families who wanted to sign up were allowed to sign up, and all of their children who had parents' permission and wanted to participate were allowed to do so. Before starting data analysis, we excluded children who had a previously diagnosed language or developmental delay, whose parents had not provided the information needed to place the child in a language or SES group, or who were not between the ages of 36 and 71 months, inclusive.

## Procedure

Testing took place at school, during school hours. Researchers first visited each classroom on two occasions, for at least an hour each time, to build rapport with the children. Testing was done at a table in a quiet room or in a hallway near the classroom. Researchers worked in pairs, with one researcher administering the task while a second researcher recorded the data using a video camera and/or pen and paper, depending on the task. Children who seemed bored or unhappy during testing were asked, "Do you want to go back to your classroom?" and returned to the classroom if they wished. All research activities were approved by the IRB of UC-Irvine [Human Subjects Protocol No. 2005-4735].

For English monolinguals, and for those bilingual learners whose home language was not Spanish, all assessments were conducted in English. For Spanish/English bilingual learners, assessments of vocabulary and symbolic number knowledge (spoken and written) were conducted twice—once in Spanish and once in English, in counterbalanced order. The assessment of nonsymbolic approximate number system (ANS) acuity was done in English, Spanish, or a combination of both, as the child preferred. Testing with Spanish/English bilingual children was carried out by fully bilingual research assistants who were native speakers of both Spanish and English. While at the schools, each of these research assistants spoke only in one language (Spanish or English) in order to give children the impression that they (the research assistants) were monolingual speakers of the language. This was done to encourage the children to cooperate with testing in the assigned language with each research assistant.

## Picture Vocabulary Test (PPVT-III)

All children were given the PPVT-III (L.M. Dunn & Dunn, 1997) to assess their English receptive vocabulary. Children were shown four pictures and asked to point to the picture that corresponded to the word spoken by the experimenter. The PPVT-III has a mean standard score of 100 and a standard deviation of 15. It was standardized on a monolingual, English-speaking sample of children from the United States and has a reported test-retest reliability of above .90 for children aged 2.5 to 6 years old (Dunn & Dunn, 1997).

## Test de Vocabulario en Imagenes Peabody (TVIP)

Spanish-speaking bilingual learners were given the TVIP (Dunn et al., 1986) to assess their receptive vocabulary in Spanish. The TVIP is similar in presentation and administration to the PPVT-III. It has a mean standard score of 100 and a standard deviation of 15. It was standardized on monolingual, Spanish-speaking samples of children from Mexico and Puerto Rico and has a reported split-half reliability of above .91 for children aged 3.0 to 6 years old. According to the authors, the test is appropriate for measuring Spanish vocabulary growth in both monolingual and bilingual children.

### Intransitive Counting

The purpose of this task was to assess children's knowledge of the verbal counting list. Children were asked to count out loud from one to ten. The experimenter introduced the task by saying, "Let's count to ten. Ready? One, two, three, four, five, six, seven, eight, nine, ten. Now you count." After the child counted once, the experimenter said, "Thanks. Can you count again?" If the child stopped counting, the experimenter prompted them to continue by repeating the last two numbers the child had said (e.g., "two, three . . . ?") or by asking "Do you know what comes after [three]?" repeating the last number the child had said. Test-retest reliability for the two verbal counts was .75.

Scores reflected the highest number the child counted to without error (maximum possible score: 10). For example, a count of "One, two, three, five, six, ten," would be scored as 3. Each child's highest score of the two counts (or four counts--two in each language for Spanish/English bilingual learners) was used. For the entailment analysis and to enable comparison with the transitive counting task, scores of 6 were coded as proficient; scores of 5 or below were coded as not proficient. (Although children were invited to count to ten, we treated counts of six and above as successes in order to facilitate comparison with the transitive counting task, where children were never asked to count rows of more than six objects.)

### Transitive Counting

In this task, children were asked to count rows of three and six items (Figure 1, Panel A). First, the experimenter placed on the table a laminated card showing a row of three, saying, "Can you show me how you count these?" If the child did not start counting immediately, the experimenter pointed to the leftmost object and said "One . . ." or "Uno . . ." and then pointed to the next object. After the row of three, the child was invited to count a row of six. If a child stopped counting before reaching the end, the experimenter repeated the last two number words the child had said with a rising intonation (e.g., "three, four...?"); or the experimenter asked, "What comes after [four]?" using the last number word produced by the child. Children were invited to count each row twice. Test-retest reliability for counts of three and six objects were .73 and .69, respectively. Counting was coded as correct up to the point of the first error, and children's highest correct count was used in the analysis (maximum possible score: 6). For the entailment analysis, scores of 6 were coded as proficient; scores of 5 or below were coded as not proficient.

### Give-N

This task, also called Give-a-Number (Wynn, 1990, 1992) was used to determine which cardinal number-word meanings the child knew (Negen et al., 2012; Sarnecka & Lee, 2009). Children's performance on this task determined their knower level (pre-number-knower, one-knower, two-knower, three-knower, four-knower, or cardinal-principle-knower, abbreviated *CP-knower*).

In this task, researchers asked children to give some number of objects between one and six to a stuffed animal. Children were asked for each number three times in a preset, randomized order for a total of 18 trials. The experimenter set up the task by placing the following items on a table: a stuffed animal (e.g., a dinosaur), a tub with small, plastic objects in it (e.g., yellow bananas), and a lid from the top of the tub, to be used as a plate (Figure 1, Panel B). The experimenter asked the child to give the dinosaur different numbers of objects (e.g., "Can you give him one banana?") by placing the objects on the dinosaur's plate and sliding it across the table to the dinosaur. After each trial, the experimenter asked a follow-up question repeating the original number, e.g., "Is that one?" Children were given generalized positive feedback (e.g., "Thank you!") on every trial, regardless of their responses. Children who responded correctly to at least two out of three requests for a given number (e.g., gave three items on at least two of the three trials where they were asked for "three") and did not give that number more than once during the 15 trials asking for other numbers were considered to "know" that number. Children's responses were sorted into number-knower levels (pre-number-knower, one-knower, two-knower, three-knower, four-knower, or CP-knower) for analysis (maximum possible score: 5). In the current sample, reliability was  $\alpha = .94$ . For the entailment analysis, CP-knowers were counted as proficient; all other number-knower levels were counted as not proficient.

## Numerical Identification

In this task, we asked children to identify the written numerals 1 through 10. Materials included a small picture frame cut from foam core and shaped like a house, with a small square of velcro in the center.

**Figure 1**

*Stimuli Used in Each of the Number Tasks*



There was also a set of laminated cards showing the written numerals zero through ten, with a square of velcro on the back of each card. The cards were placed face up on the table in an unordered jumble (Figure 1, Panel C.). The experimenter introduced the game by saying, “Here are some numbers. I’m going to look for zero. This is zero.” (Here the experimenter picked up the laminated card with the number zero on it and showed it to the child.) “I’m going to put zero in its home, like this.” (Here the experimenter placed the number card in the house-shaped frame sticking the velcro squares together, and then pulled the card off the house and returned it to the pile.) “Now it’s your turn. Can you find the number *one* and put it in its home?” When the child placed a number card in the frame, the experimenter asked a follow-up question repeating the original number (e.g., “Is that one?”) Children were asked for the numbers one through



ten in ascending order and were given generalized positive feedback (e.g., “Thank you!”) on every trial, regardless of their responses. Each trial was coded as correct or incorrect, regardless of orientation (maximum possible score: 10). For example, if a child placed the correct number in the frame, but placed it there upside-down, it was still counted correct. The exceptions to this rule were the numbers 6 and 9, which had to be placed right-side-up in order to count as correct. In the current sample, reliability was  $\alpha = .92$ . For the entailment analysis, scores of 10 were coded as proficient; scores of 9 or below were coded as not proficient.

### Scaffolded Number Line

This task was developed as a simpler alternative to the commonly-used continuous number line task (e.g., Ramani & Siegler, 2008). Its purpose was to assess children’s understanding that the written number line represents numbers in ascending order from left to right. Materials included a rectangular piece of foam core with a line of eleven velcro squares on it, and the laminated cards from the previous task (Figure 1, Panel D). The experimenter set up the task by placing the numbers zero and ten on the first and last squares respectively, saying “This is a number line. It starts with zero and goes to ten. And, this is five.” Here the experimenter picked up the card with the number five on it. “It goes right in the middle, like this.” Here the experimenter placed the five on the fifth velcro square of the board. “Now it’s your turn. I’ll give you a number and you put it in its place on the number line. Okay, here’s the number seven. Can you show me where it goes on the number line?”

Children were asked for the numbers one through four and six through nine in a preset, randomized order (maximum possible score: 8). They were given generalized positive feedback (e.g., “Thank you!”) on every trial, regardless of their responses. After each trial, the experimenter removed the number from where the child had placed it on the number line (prior to asking for the next number). Children’s responses were coded as correct or incorrect. In the current sample, reliability was  $\alpha = .90$ . For the entailment analysis, scores of 8 were coded as proficient; scores of 7 or below were coded as not proficient.

### Continuous Number Line

In this task, the experimenter placed a strip of paper on the table in front of the child. On the paper was printed a horizontal line, with the numeral 0 at the left end and the numeral 10 at the right end (Figure 1, Panel E, adapted from Ramani & Siegler, 2008). The experimenter said to the child, “Now we’re going to play another game with a number line. Just like the other number line, this number line starts with zero and goes to ten. In this game, you *point* to where the number should go. OK? So if I say, ‘Where does FIVE go,’ you would put your finger right in the middle of the line, like this, because FIVE is in the middle, between zero and ten.” The child was guided to place their finger on the line, and then the experimenter used a pencil to make a tick mark on the line where the child’s finger was. The child’s score was calculated as the correlation between each tick mark and the correct location on the line, with scores ranging between -1 and 1 (maximum possible score: 1). Higher scores represented better performance. In the current sample, reliability was  $\alpha = .73$ . For the entailment analysis, scores of .96 or above were coded as proficient; scores of 0.95 or below were coded as not proficient.

### Nonsymbolic Number Comparison (ANS Acuity)

For each trial of the task, children were shown a laminated card, 21.5cm long by 12.5cm high, with two clouds of black dots printed side by side (Figure 1, Panel F). Each cloud contained 20-100 dots. The experimenter introduced the task by saying, “Look at this card. This card has two sides. There are some dots on this side, and some dots on this side. (Experimenter points to dot clouds on the left and right.) You need to point to the side that has more dots. Which side has more dots?”

First, children were given training trials to make sure that they understood the task (Negen & Sarnecka, 2015). The ratio of the dot clouds in these training trials was 1:3 (e.g., one cloud had 30 dots while the other had 90). Even preverbal infants can discriminate numbers at a ratio of 1:3 (Feigenson et al., 2004; Izard et al., 2009). On the training trials, the experimenter gave children feedback directing their attention to the number of dots. For example, if a child chose the side with fewer but bigger dots, the experimenter said: “Well these dots are *bigger*, but this side has *more* dots.”

They're *smaller*, but there's *more* of them." Children were presented with training trials until they either answered eight training trials in a row correctly or chose to stop playing.

After answering correctly on eight training trials in a row, children were given test trials. These consisted of nine blocks with ratios at 1:2 (= .50), 7:12 (= .58), 2:3 (= .66), 17:24 (= .71), 3:4 (= .75), 4:5 (= .80), 5:6 (= .83), 7:8 (= .87), and 9:10 (= .90). There were eight trials per block, and all trials within each block contained the same ratio. Feedback was given after every trial (e.g., "That's right— this side has more." or "Uh oh, this side has more dots, you see?"). Trials were presented too rapidly for children to count the dots, and no children were observed attempting to count. Children's performance was scored as the average of the number of correct trials in each block (maximum possible score: 8). In the current sample, reliability was  $\alpha = .62$ .

## Results

### Demographic Variables

Of the 513 children who contributed data to the analysis, 77% ( $n = 393$ ) were bilingual or multilingual learners and 23% ( $n = 120$ ) were monolingual English learners. Among the bilingual/multilingual learners, the most common home language was Spanish (65%,  $n = 254$ ), followed by Chinese (all forms, 16%,  $n = 61$ ). The remaining 78 children spoke 29 different home languages, with no single language accounting for more than 3% ( $n = 13$ ) of participants.

The majority of parents (92%,  $n = 470$ ) reported their highest educational attainment. Among these parents, parental education ranged from: 1st-11th grade 23% ( $n = 109$ ; median highest grade completed was 8th grade); completed high school or GED: 22% ( $n = 103$ ); two years of college or Associate's degree: 11% ( $n = 51$ ); four years of college or Bachelor's degree: 14% ( $n = 67$ ); Master's or Doctoral degree: 30% ( $n = 140$ ). The majority of parents (92%,  $n = 473$ ) also reported their annual household income. Reported incomes ranged from \$0 - \$10,000 annually to over \$75,000 annually. The median income range was \$20,000 - \$30,000.

### Research Question 1. Relations Among Numeracy Skills

Descriptive statistics for each numeracy skill are presented in Table 1. Bivariate correlations between each numeracy skill and parental education, household income, degree of bilingualism (0 = 0-5% non-English in the home; 1 = 5-50% non-English in the home; 2 = 50-95% non-English in the home; 3 = 95-100% non-English in the home), and children's English receptive vocabulary (PPVT) scores show that all of the symbolic numeracy skills were positively correlated (Figure 2) with correlation coefficients ranging from moderate to strong effects (e.g.,  $r_s = .20 - .79$ ). Indicators of socioeconomic status (parental education and household income) were more strongly correlated to symbolic numeracy skills than degree of bilingualism (for more detailed analyses, see Research Question 2, below).

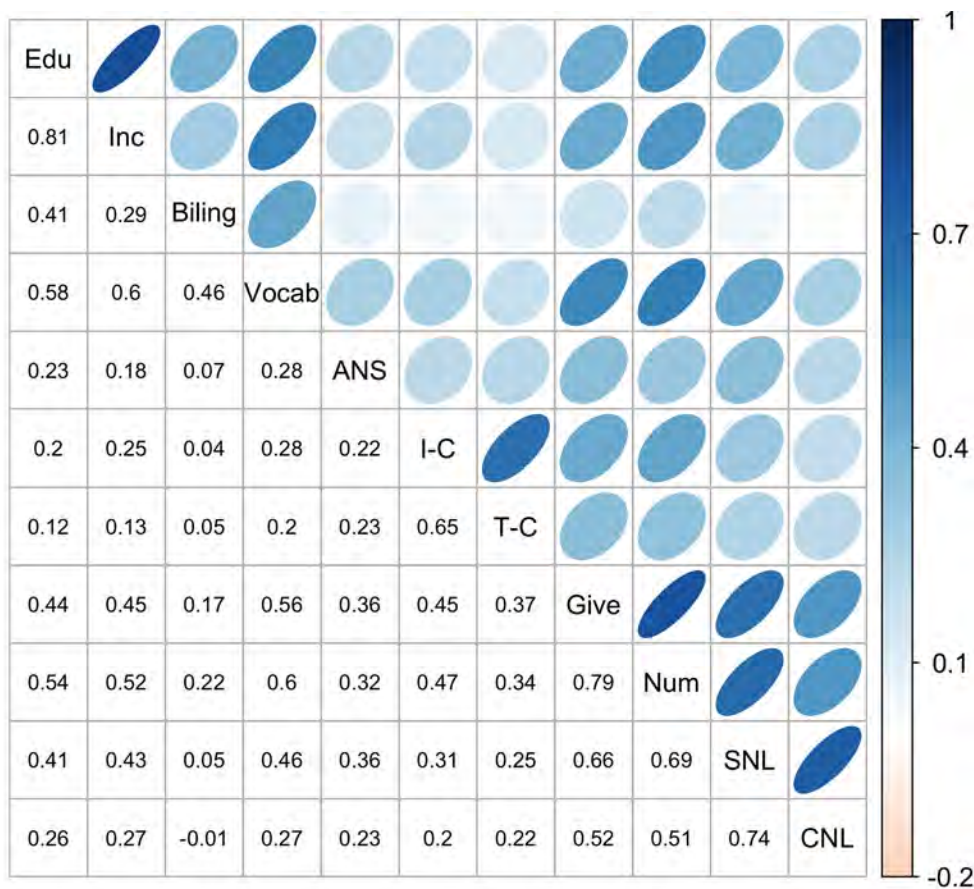
**Table 1**

*Descriptive Statistics of Children's Performance on Vocabulary and Numeracy Tasks*

Task	<i>N</i>	<i>M</i>	<i>SD</i>	Minimum	Maximum
Vocabulary	502	89.71	19.49	40.00	130.00
Nonsymbolic number comparison	358	5.31	0.67	3.33	7.11
Intransitive counting	510	9.17	2.07	1.00	10.00
Transitive counting	512	5.68	0.97	1.00	6.00
Give-N	510	3.43	1.70	0.00	5.00
Number identification	490	6.35	3.64	0.00	10.00
Scaffolded number line	477	3.56	2.89	0.00	8.00
Continuous number line	495	0.47	0.47	-0.98	1.00

Figure 2

Bivariate Correlations Between Demographic Variables and Numeracy Skills



Note. The shapes in the upper right portion of the figure show the same information as the text in the lower left portion, with stronger correlations represented by darker, narrower ovals. Correlations of .12 and .13 (Edu and T-C; Inc and T-C)  $ps < .01$ ; correlations less than .08 (Biling and ANS, I-C, T-C, SNL, CNL)  $ps > .05$ ; all other correlations  $ps < .001$ . Edu stands for “education”; Inc stands for “income”; “Biling” stands for “bilingualism”; Vocab stands for “vocabulary”; ANS stands for “approximate number system” and represents the nonsymbolic number comparison task; I-C stands for “intransitive counting”; T-C stands for “transitive counting”; Give stands for “Give-N”; Num stands for “numeral identification”; SNL stands for “scaffolded number line”; CNL stands for “continuous number line.”

### Nonsymbolic Number Comparison

The changes that we made to the ANS testing procedure ensured that for those children who did complete the task (that is, who reached criterion on the training trials and then went on to complete at least one block of test trials), we can be confident that what we measured really was their ANS acuity. However, these changes also meant that a lot of children simply could not complete the task. Nearly one-third of the children we tested ( $n = 155$ ) never succeeded at eight training trials in a row, and so never moved on to the test trials. A higher proportion of children who passed the ANS training were also CP-knowers (i.e., earned a score of 5 on the Give-N task; 58.4%) than children who did not pass the ANS training (17.4%).

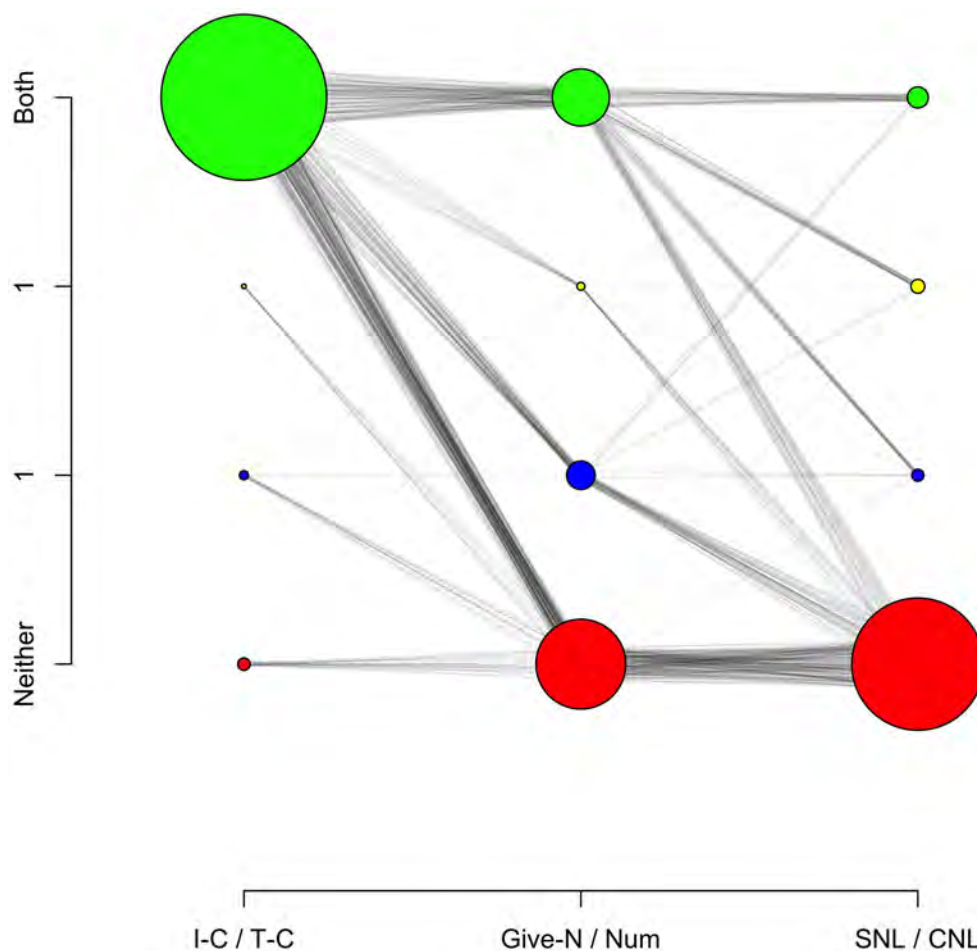
### Entailment Relations Among Symbolic Skills

The entailment relations among the early symbolic numeracy skills are illustrated in a parallel plot (Figure 3). This allows us to visualize which symbolic numeracy skills may entail (or build on) other skills. First, children were categorized as proficient or not proficient on each task (see Method section). The size of each circle represents the relative number of children who were categorized as proficient vs. not proficient on each pair of tasks (intransitive and

transitive counting; Give-N and numeral identification; scaffolded and continuous number line). Each line represents a single child's performance on one pair of skills relative to the next pair, with the overall density of lines showing the most common patterns of children's proficiency across tasks. The pattern shows that proficiency on each pair of tasks is necessary (but not sufficient) for children to succeed on the next pair of tasks. For example, most children were proficient on both the intransitive and transitive counting tasks (large, green circle at top left). Most children were not proficient at either the scaffolded or continuous number line tasks (large, red circle at bottom right). Proficiency on both the intransitive and transitive counting tasks appeared to be necessary, but not sufficient, for success on the Give-N and Numeral ID tasks. Proficiency on Give-N and Numeral ID was, in turn, necessary but not sufficient for success on the number-line tasks. The relatively few lines sloping upward from left to right represent individual children who failed at one of the earlier tasks and succeeded at a later one, underlining the finding that the earlier skills seemed to be prerequisites for the later ones.

**Figure 3**

*Entailment Relations Among Symbolic Number Skills*



*Note.* I-C stands for “intransitive counting”; T-C stands for “transitive counting”; Num stands for “numeral identification”; SNL stands for “scaffolded number line”; CNL stands for “continuous number line”.

## Research Question 2. Numeracy Skills and Bilingualism/Monolingualism

We examined the relation between symbolic number skills (those using spoken or written numerals) and language use to look for differences in patterns of development for bilingual vs. monolingual learners, separating these from other well-documented socio-economic influences such as household income and parental education. For each symbolic

number skill, we present the results of a traditional frequentist multiple linear regression with bilingualism and household income as predictors, controlling for child age. For this analysis, language context was dichotomized as bilingual vs. monolingual, with households using English at least 95% of the time considered monolingual, and all others considered bilingual. (All children in the study attended English-speaking preschool programs, so there were no monolingual learners of languages other than English.) We calculated the multiple linear regression models using the *stats* package in RStudio version 1.2.1335 (RStudio Team, 2020). We also report a Bayes-factor model comparison for predicting each symbolic number skill. The Bayes factor model comparisons are each based on three models: one that included bilingualism, education, income, and child age; one that included education, income, child age and not bilingualism; and one that included just an intercept and no other factors. This allows us to compare the predictive performance of competing models and summarize the plausibility of each variable in the model with a Bayes Factor (BF, a number that summarizes how much more plausible the predictive model is given that variable; (Morey & Rouder, 2011; Rouder & Morey, 2012). We conducted all Bayesian linear regression models using the *BayesFactor* package in RStudio (Morey & Rouder, 2018).

### Intransitive Counting

Bilingualism was not a significant predictor of children's intransitive counting skills after controlling for parental education, household income, and child age ( $B = 0.27$ ,  $SE = 0.23$ ,  $t = 1.13$ ,  $p = .259$ ). Age was a significant predictor of children's intransitive counting skills,  $B = 0.05$ ,  $SE = 0.01$ ,  $t = 3.51$ ,  $p < .001$ . Parental education was not a significant predictor children's intransitive counting skills,  $B = 0.12$ ,  $SE = 0.10$ ,  $t = 1.18$ ,  $p = .238$ , however, household income was significantly related to children's intransitive counting skills,  $B = 0.14$ ,  $SE = 0.14$ ,  $t = 2.43$ ,  $p = .016$ . Across monolinguals and bilinguals with varying levels of parental education, children from higher income households tended to reach higher numbers when counting than children from lower income households (average of count 8.8 for household incomes below \$30,000 per year versus average count of 9.5 for household incomes above \$30,000 per year). Similarly, the winning Bayesian model was that with parental education, income, and child age without bilingualism (Table 2). Including bilingualism lowered the Bayes factor by a factor of 2.89.

**Table 2**

*Bayes Factor Model Comparison*

Task	Intercept-only model	Age, education, income model	Age, education, income, bilingualism model
Intransitive counting	1 to 571167	*	1 to 2.89
Transitive counting	1 to 502	*	1 to 3.48
Give-N	1 to 4.66	*	1 to 6.46
Numeral identification	1 to 2.04	*	1 to 5.50
Scaffolded number line	1 to 10 <sup>47</sup>	1 to 1.19	*
Continuous number line	1 to 10 <sup>26</sup>	*	1 to 1.45

*Note.* Asterisks mark the preferred model for each numeracy skill. The remaining values are the Bayes factors between a model and the preferred model for each skill.

### Transitive Counting

There was not a significant effect of bilingualism on children's transitive counting skills after controlling for parental education, household income, and child age,  $B = 0.10$ ,  $SE = 0.11$ ,  $t = 0.86$ ,  $p = .393$ . Age was a significant predictor of children's transitive counting skills,  $B = 0.03$ ,  $SE = 0.01$ ,  $t = 4.16$ ,  $p < .001$ . Parental education was not significantly related to children's transitive counting skills,  $B = 0.08$ ,  $SE = 0.05$ ,  $t = 1.61$ ,  $p = .107$ , nor was household income,  $B = 0.01$ ,  $SE = 0.03$ ,  $t = 0.47$ ,  $p = .637$ . The winning Bayesian model was that with parental education, income, and child age without bilingualism (Table 2). Including bilingualism lowered the Bayes factor by a factor of 3.48.

### Give-N

There was not a significant effect of bilingualism on children's Give-N performance after controlling for parental education, household income, and child age,  $B = 0.14$ ,  $SE = 0.16$ ,  $t = 0.91$ ,  $p = .362$ . Age was a significant predictor of children's Give-N performance,  $B = 0.11$ ,  $SE = 0.01$ ,  $t = 11.40$ ,  $p < .001$ . Parental education was a significant predictor of children's Give-N performance,  $B = 0.30$ ,  $SE = 0.07$ ,  $t = 4.39$ ,  $p < .001$ , as was household income,  $B = 0.14$ ,  $SE = 0.04$ ,  $t = 3.75$ ,  $p < .001$ . Children whose parents had higher levels of education tended to have higher knower-levels than children whose parents had lower levels of educational attainment (average knower-level of 4.0 for children whose parents completed some college or more versus average knower-level of 2.7 for children whose parents completed high school or fewer years of education). Children from higher income households tended to have higher knower-levels than children from lower income households (average knower-level of 4.3 for household incomes above \$30,000 per year versus average knower-level of 2.7 for household incomes below \$30,000 per year). The winning Bayesian model was that with parental education, income, and child age without bilingualism (Table 2). Including bilingualism lowered the Bayes factor by a factor of 6.46.

### Numerical Identification

There was not a significant effect of bilingualism on children's numeral identification skills after controlling for parental education, household income, and child age,  $B = 0.37$ ,  $SE = 0.34$ ,  $t = 1.11$ ,  $p = .269$ . Age was a significant predictor of children's numeral identification skills,  $B = 0.18$ ,  $SE = 0.02$ ,  $t = 8.85$ ,  $p < .001$ . Parental education was significantly related to children's numeral identification skills,  $B = 0.88$ ,  $SE = 0.15$ ,  $t = 5.83$ ,  $p < .001$ , as was household income,  $B = 0.31$ ,  $SE = 0.08$ ,  $t = 3.67$ ,  $p < .001$ . Children whose parents had higher levels of education tended to correctly identify more numbers than children whose parents had lower levels of educational attainment (average score of 7.9 for children whose parents completed some college or more versus average score of 4.4 for children whose parents completed high school or fewer years of education). Children from higher income households tended to correctly identify more numbers than children from lower income households (average score of 8.5 for household incomes above \$30,000 per year versus average score of 4.6 for household incomes below \$30,000 per year). The winning Bayesian model was that with parental education, income, and child age without bilingualism (Table 2). Including bilingualism lowered the Bayes factor by a factor of 5.50.

### Scaffolded Number Line

Bilingualism was a significant predictor of children's scaffolded number line performance after controlling for parental education, household income, and child age,  $B = 0.56$ ,  $SE = 0.27$ ,  $t = 2.09$ ,  $p = .037$ . Controlling for parental education, household income and child age, bilingual children tended to be more accurate in their scaffolded number line performance than monolingual children (marginal mean of bilingual children of 3.66,  $SE = 0.12$  versus marginal mean of monolingual children of 3.10,  $SE = 0.23$ ). Age was a significant predictor of children's scaffolded number line performance,  $B = 0.20$ ,  $SE = 0.02$ ,  $t = 12.47$ ,  $p < .001$ . Parental education was significantly related to children's scaffolded number line,  $B = 0.38$ ,  $SE = 0.12$ ,  $t = 3.27$ ,  $p = .001$ , as was household income,  $B = 0.36$ ,  $SE = 0.07$ ,  $t = 5.47$ ,  $p < .001$ . Children whose parents had higher levels of education tended to have higher scores than children whose parents had lower levels of educational attainment (average score of 4.99 for children whose parents completed some college or more versus average score of 2.43 for children whose parents completed high school or fewer years of education). Children from higher income households tended to be more accurate on the scaffolded number line than children from lower income households (average score of 5.0 for household incomes above \$30,000 per year versus average score of 2.4 for household incomes below \$30,000 per year). The winning Bayesian model was that with bilingualism, parental education, income, and child age (Table 2). Removing bilingualism lowered the Bayes factor by a factor of 1.19.

### Continuous Number Line

Bilingualism was not a significant predictor of children's continuous number line performance after controlling for parental education, household income, and age,  $B = 0.09$ ,  $SE = 0.05$ ,  $t = 1.86$ ,  $p = .063$ . Age was a significant predictor of children's continuous number line,  $B = 0.03$ ,  $SE = 0.003$ ,  $t = 10.49$ ,  $p < .001$ . Parental education significantly predicted children's continuous number line performance,  $B = 0.05$ ,  $SE = 0.2$ ,  $t = 2.35$ ,  $p = .019$ , as did household income,  $B = 0.03$ ,

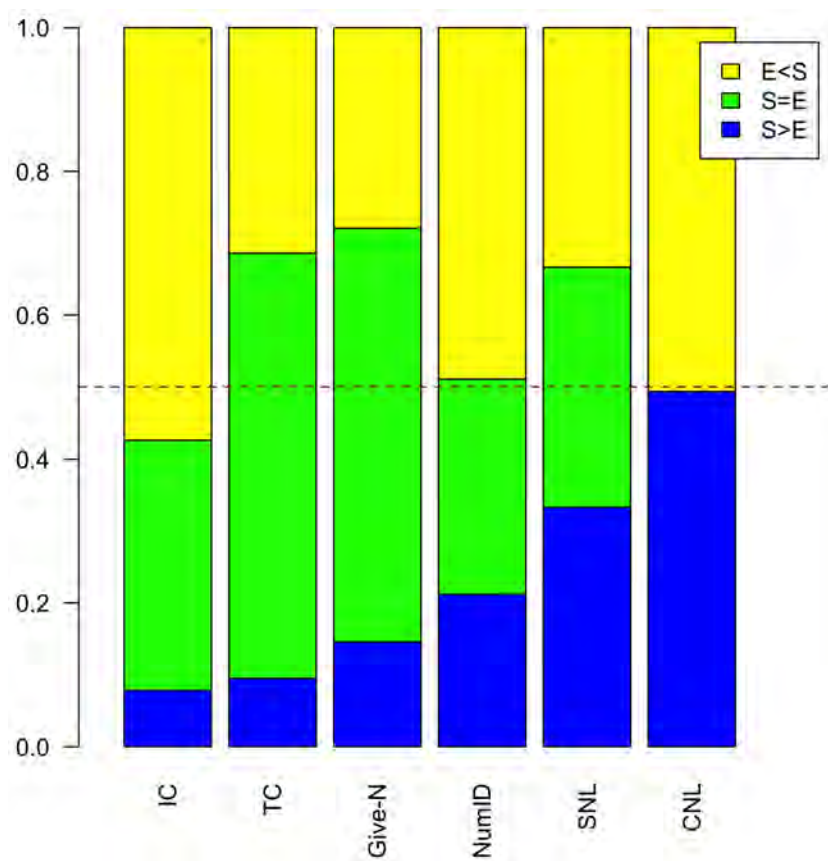
$SE = 0.01$ ,  $t = 2.81$ ,  $p = .005$ . Children whose parents had higher levels of education tended to have higher scores than children whose parents had lower levels of educational attainment (average score of 0.55 for children whose parents completed some college or more versus average score of 0.35 for children whose parents completed high school or fewer years of education). Children from higher income households tended to be more accurate on the continuous number line than children from lower income households (average score of 0.62 for household incomes above \$30,000 per year versus average score of 0.35 for household incomes below \$30,000 per year). The winning Bayesian model was that with parental education, income, and child age without bilingualism (Table 2). Including bilingualism lowered the Bayes factor by a factor of 1.45.

### Research Question 3. Numeracy in Home Language Versus Instructional Language

In our third research question, we were interested in whether and how bilingual children’s symbolic number knowledge differed when tested in their home language vs. the language of their preschool instruction. To address this question, we considered only the Spanish/English bilingual learners in our sample ( $N = 254$ ; 94% attended Head Start programs), comparing their performance on the English and Spanish versions of each task (see Figure 4). Below we report the results of frequentist paired samples  $t$ -tests as well as a Bayesian version using the *BayesFactor* package in R (Morey & Rouder, 2018), which performs the JZS  $t$ -test as described in Rouder et al. (2009).

**Figure 4**

*Comparison of Spanish-English Bilingual Learners’ Performance on Each Task When Administered in English vs. Spanish*



*Note.* Blue areas show the proportion of children who scored higher when tested in Spanish than in English; green areas show the proportion who scored the same in both languages; yellow areas show the proportion who scored higher in English. In the figure caption, E stands for “English” and S stands for “Spanish”. On the axis, IC stands for “intransitive counting”; TC stands for “transitive counting”; NumID stands for “numeral identification”; SNL stands for “scaffolded number line”; CNL stands for “continuous number line”.

## Vocabulary Measures

The PPVT and TVIP were used to measure children's receptive vocabulary in English and Spanish, respectively. On the PPVT, the mean standard score was 78.50 ( $SD = 17.12$ ). On the TVIP, the mean standard score was 81.64 ( $SD = 15.79$ ). Both of these mean standard scores fell outside the normal range—more than one standard deviation ( $SD = 15$ ) away from the population mean ( $\mu = 100$ ). Spanish-speaking bilinguals scored significantly better on the TVIP than on the PPVT,  $t(266) = -4.30$ ,  $p = .007$ . The Bayesian  $t$ -test supported the alternative hypothesis by a factor of 2.89, providing additional evidence that the two means were different.

## Intransitive Counting

This task measured children's ability to correctly count out loud to ten; thus, the maximum score possible was 10. The average counting out loud score was 6.22 ( $SD = 3.24$ ) in Spanish and 8.66 ( $SD = 2.66$ ) in English. Children performed significantly better in English than Spanish,  $t(241) = 12.05$ ,  $p < .001$ . While only 32% of children performed at ceiling (i.e., counted out loud correctly up to 10) in Spanish, 75% did so in English. The Bayesian  $t$ -test supported the alternative hypothesis by a factor of 2.00, providing additional evidence that the two means were different.

## Transitive Counting

This task measured children's ability to correctly count six objects; thus, the maximum score possible was 6. In Spanish, the mean score was 4.80 ( $SD = 1.82$ ). In English, the mean score was 5.47 ( $SD = 1.22$ ). Children performed significantly better in English than Spanish,  $t(122) = 5.70$ ,  $p < .001$ . However, most children performed at ceiling (i.e., counted the six objects correctly) in both languages: 61% of children performed at ceiling in Spanish and 82% did so in English. The Bayesian  $t$ -test supported the alternative hypothesis by a factor of  $2.5 \times 10^5$ , providing additional evidence that the two means were different.

## Give-N

This task measured children's knowledge of the exact meanings of the numbers "one" through "six," to determine their knower level in each of their languages; the highest knower level possible was CP-knower. When assessed in Spanish, there were 28 pre-knowers (12%), 50 one-knowers (21%), 68 two-knowers (28%), 47 three-knowers (19%), 19 four-knowers (8%), and 30 CP-knowers (12%). When assessed in English, there were 18 pre-knowers (7%), 62 one-knowers (25%), 51 two-knowers (20%), 44 three-knowers (18%), 16 four-knowers (6%), and 60 CP-knowers (24%). Children performed significantly better in English than Spanish,  $t(239) = 4.19$ ,  $p < .001$ . There was a strong positive correlation between children's knower-level in Spanish and English,  $r(238) = .788$ ,  $p < .001$ ; 58% of children performed equally in both languages, 28% performed better in English, and 15% performed better in Spanish. The Bayesian  $t$ -test supported the alternative hypothesis by a factor of 316, providing additional evidence that the two means were different.

## Numeral Identification

This task measured children's ability to correctly identify the written numerals one through ten; thus, the maximum score possible was 10. The average numeral identification score was 3.41 ( $SD = 2.93$ ) in Spanish and 4.31 ( $SD = 3.61$ ) in English. Children performed significantly better in English than Spanish,  $t(230) = 6.33$ ,  $p < .001$ . However, the most common score in both languages was a 1, with 20% of children receiving that score in Spanish and 20% of children receiving that score in English. This suggests that a cluster of children appear to simply be guessing (one would expect a score of 1 on average from guessing). Yet, 14% of children performed at ceiling in English and only 5% did so in Spanish. The Bayesian  $t$ -test supported the alternative hypothesis by a factor of  $6.1 \times 10^6$ , providing additional evidence that the two means were different.

## Scaffolded Number Line

This task measured children's ability to correctly place the numbers one through nine (excluding five) on a number line; thus, the maximum score possible was 8. In Spanish, the mean score was 1.84 ( $SD = 2.16$ ). In English, the mean score was 1.98 ( $SD = 2.22$ ). Children's performance in English versus Spanish did not differ significantly,  $t(253) = 1.46$ ,  $p = .146$ . As



with the Numeral Identification task, a significant portion of children appear to be guessing; the most common score in both languages was a 1 (one would expect a score of 1 on average from guessing), with 30% of children receiving that score in Spanish and 32% of children receiving that score in English. The Bayesian  $t$ -test resulted in a Bayes factor for the alternative hypothesis by a factor of 0.20, providing additional evidence that the two means were not different.

### Continuous Number Line

This task measured children's ability to correctly place the numbers one through nine (excluding five) on a number line, with scores calculated as the correlation between each placement mark and the correct location on the line. Thus, the maximum possible score was 1 (perfect correlation). In Spanish, the mean score was 0.20 ( $SD = 0.45$ ). In English, the mean score was also 0.20 ( $SD = 0.48$ ). Children's performance in English versus Spanish did not differ significantly,  $t(234) = 0.18$ ,  $p = .855$ . The Bayesian  $t$ -test resulted in a Bayes factor for the alternative hypothesis by a factor of 0.07, providing additional evidence that the two means were not different.

## Discussion

This study investigated early childhood number skills in a linguistically and economically diverse sample of preschoolers in Orange County. The study had three goals: To describe entailment relations among a set of early numeracy skills; to examine the relation between symbolic number-skill development and bilingualism/monolingualism when controlling for known effects of socioeconomic status; and to find out whether and how bilingual learners' symbolic number knowledge differed when tested in their home language vs. the language of preschool instruction. To these ends, children were tested in English (and also in Spanish, where appropriate) on a battery of tasks related to symbolic number skills, nonsymbolic number comparison, and receptive vocabulary.

The first set of questions related to entailment relations among early numeracy skills. Logic dictates that certain skills build on others. For example, intransitive counting (reciting the number list) is part of transitive counting (reciting the number list and pointing to objects); and both of these provide a foundation for children to use counting to construct sets on the Give-N task. But not all relations among tasks are so easy to predict. We were interested both in symbolic skills (those using spoken and written numerals) and in nonsymbolic number comparison, which we tested by asking children to decide which of two cards had "more dots."

Among the symbolic skills, there was strong evidence that numeracy skills build on each other, in mostly unsurprising ways. Children learn to count, then learn how counting represents number, as demonstrated by the fact that transitive and intransitive counting skills were jointly entailed by the Give-N task. Similarly, it is not surprising that children who can place number symbols correctly on a number line can also identify those written symbols. These findings replicate previous work showing that children learn the counting procedure before they master the cardinality principle (Briars & Siegler, 1984; Fuson, 1992; Sarnecka & Lee, 2009; Wynn, 1992), and that children typically have higher accuracy on numeral identification tasks than on number line tasks (Kolkman et al., 2013; Ramani & Siegler, 2008; Sasanguie et al., 2012)

Perhaps more surprising was the link between the Give-N task (where children used counting to construct sets of a given size) and the Numeral ID task (where children identified the written number symbol matching a given number word). We had no *a priori* reason to expect performance on the Give-N and Numeral ID tasks to be closely related. In principle, a child's understanding of the cardinal principle (measured by success on the Give-N task) might be quite separate from their knowledge of written numerals. Indeed, previous research has shown that 4-year-old children's cardinal knowledge of number words is significantly greater than their cardinal knowledge of written numerals and more accurate than their verbal labels of numerals (Knudsen et al., 2015). We were surprised to see that they were very highly correlated in this sample, such that in general only children who had mastered both Give-N and Numeral ID successfully navigated the number-line tasks. It appears that both cardinal-principle understanding and familiarity with written numerals are used in making sense of a number line, where numerals are arranged according to their cardinal values. One implication of this work is that interventions using number lines are unlikely to be helpful until children understand counting and cardinality, and have linked spoken number words to written number symbols.

Unlike the clear entailments found among the symbolic number tasks, we found little correlation between children's performance on a nonsymbolic number comparison task (testing approximate number system, or ANS, acuity) and their performance on any of the symbolic number tasks. Methods for assessing ANS acuity in preschoolers must be chosen carefully, as many children may not interpret the phrase "more dots" to refer to numerosity until they have a grasp of the cardinality principle (Negen & Sarnecka, 2015). In the present study, we designed the testing procedure to ensure that we were in fact measuring ANS acuity, which meant excluding all data from children whose performance did not reflect ANS judgments at very easy-to-compare ratios of 1:3. In the end, we had to exclude data from nearly one-third of the children in our sample. Of the children who did interpret the phrase "more dots" in a numerical sense, more than half understood the cardinality principle as measured by the Give-N task, whereas there were few cardinality-principle-knowers among the children whose ANS data were excluded. In the end, we found little correlation between performance on the nonsymbolic task and any of the symbolic tasks; the most important correlation was the one between being a cardinality-principle knower and understanding the ANS task instructions at all.

A second question that we sought to answer with this study was how the child's language environment (specifically comparing monolingual to bilingual or multilingual environments) relates to the early development of number skills. We wanted to use this large and diverse sample to separate bilingualism from well-documented socio-economic influences such as household income and parental education. In terms of demographics, we found abundant and unsurprising evidence that economic hardship is bad for early learning: Children whose families were under the greatest economic stress showed the most delayed acquisition of early numeracy. This is obviously of concern, given that children who enter kindergarten with poorer understanding of numbers often continue to struggle throughout the early years of elementary school (Duncan et al., 2007; Jordan et al., 2009; Watts et al., 2014). However, once we disentangled language environment from socio-economic status, data from the many bilingual learners in our sample showed no relation between the amount of English spoken in a young child's home and that child's early math learning. Contrary to supposition that bilingual learners might face challenges in early number learning because of a possible mismatch between their home language and their preschool language, we found that bilinguals were not delayed relative to monolingual English learners. And contrary to supposition that bilingualism might confer an advantage in early number learning (Bialystok & Codd, 1997), we found little evidence of a bilingual advantage either. (Bilingual learners did perform very slightly better on the scaffolded number line task, but the Bayesian evidence was so weak as to be considered anecdotal.) This suggests that bilingualism was not very important for explaining children's performance on any of the symbolic number tasks when we use rigorous controls for socioeconomic status (e.g., income and parent education). One caveat is that due to ceiling effects on the intransitive and transitive counting tasks, we weren't able to fully address whether or not bilingualism explains variability in children's performance above and beyond parent education and household income. Future research could test this more conclusively using counting tasks with more variability, or perhaps focusing exclusively on younger participants earlier on in their numerical development.

A third goal of the present study, concerning just the Spanish/English bilingual learners in our sample, was to compare children's performance on symbolic number tasks when tested in Spanish versus when tested in English. English was the language of instruction in their preschool programs; Spanish was the language used all or part of the time at home. Assessments of the children's receptive vocabulary in each language showed that most had larger vocabularies overall in Spanish than in English. However, when it came to counting, constructing sets on the Give-N task, and identifying written numerals on the Numeral ID task, most children performed as well or better in English--the language of instruction at their preschools--than they did in Spanish. We take this as evidence that preschool programs had provided these children with important numeracy-building experiences in English that went beyond the experiences that their home environments provided in Spanish. Even though the at-school experiences took place in English, which was not the most familiar language for most of these children, those experiences were still key to their developing numeracy skills. Unlike older bilingual learners for whom the challenge may be "translating" prior knowledge from one language into another, preschoolers are likely to be learning skills for the first time. And because children under five essentially become native speakers of whatever language they are exposed to, experiences with counting and number in any language are valuable. On a policy level, these findings would seem to argue for providing preschool services to as many children as possible (even if such services can only be provided in English) rather than providing fewer children with bilingual services, which may be more expensive to implement. Limited resources for bilingual services should be

allocated first to older students, who have more prior knowledge in their home language, and who face longer delays in reaching the level of English proficiency needed for instruction in English at higher grades.

The fact that the Spanish/English bilingual learners in our sample performed as well or better in English than in Spanish also has methodological implications for researchers. Testing children twice (once in each language) requires nearly twice the effort and expense of testing them only once. The present study, for example, employed native Spanish-speaking research assistants who spoke to the children only in Spanish and pretended not to understand English at all, as well as English-speaking research assistants who interacted with the children only in English. But in the end, the scientific value added by doing this was minimal: The picture of children's numeracy that we gleaned from testing separately in Spanish and English was not much different from the picture that we would have gotten by testing only in English. The results may well have been different with older students, but we hope this knowledge will be useful to our colleagues as they plan studies with very young bilingual learners in the future.

## Limitations

There are several limitations to the present study. First, these data present the results of concurrent associations among children's skills, as opposed to longitudinal or cross-sectional analyses. Although the entailment analyses suggest that some skills are mastered by children earlier than others, further research is needed to confirm directionality and potential causal connections among these early numeracy skills. Second, these data do not provide an exhaustive comparison of all early numeracy skills. In particular, future research should test whether these patterns are observed with nonsymbolic comparisons of additional set sizes and symbolic comparisons of pairs of numerals (e.g., which is more, 3 or 4?). Third, our sample includes preschoolers ranging in age (3 - 6 years old) and prior years of preschool (e.g., enrolled in their first or second year of Head Start). It remains an open question as to whether the pattern of results vary based on children's age or exposure to the language of instruction—for instance, whether children's comparable performance on measures of number skills is true of the youngest learners with the least exposure to English in the classroom setting.

## Conclusions

Overall, our findings with this large, linguistically and socioeconomically diverse sample confirm that early number skills such as counting out loud, constructing sets, and recognizing written numerals form an important base for later mathematics learning. Even the preschool-level tasks included in this study build on each other, with earlier-acquired skills such as counting serving as prerequisites for later skills such as using a number line. Nonsymbolic estimation acuity, on the other hand, seems less important in these early years. Furthermore, unlike some other aspects of language development, early numeracy is hardly affected by a child's being monolingual or bilingual/multilingual. More important are the preschool experiences that support children's growing number concepts, whether or not those experiences occur in the child's home language.

Although this study included a relatively large and diverse sample of participants, the developmental variation among the preschoolers of Orange County does not equal the variation in young children across the U.S. nation and beyond. We hope that this study will be part of a growing, more inclusive literature on development in early childhood. The more we as a field reach beyond convenience samples, the more our scholarship can help educators and policymakers support the growth and well-being of all our children.

---

**Funding:** This research was supported by NSF grant DRL-0953521 to the first author. Any opinions, findings, conclusions, or recommendations expressed here are those of the authors, and do not necessarily reflect the views of the National Science Foundation.

---

**Acknowledgments:** The authors would like to thank the children, families, preschool teachers and administrators who participated in this research, as well as lab manager Tanya Anaya, Prof. Dr. Lieven Verschaffel, and two anonymous reviewers for their help with this project.

---

**Competing Interests:** The authors have declared that no competing interests exist.

---

**Data Availability:** For this article, a data set is freely available (Sarnecka, Negen, Scalise, Goldman, & Rouder, 2022).

---

## Supplementary Materials

The Supplementary Materials contain the following items (for access see [Index of Supplementary Materials](#) below):

- Research data
- Stimuli and data collection sheets
- Participants' household incomes by preschool type
- Descriptive statistics on vocabulary and numeracy measures, by child age
- Floor and ceiling effects observed on numeracy tasks
- Multiple linear regression models for each numerical skill and the amount of variance explained by each model
- Entailment relations among symbolic number skills, with a child's performance coded as "proficient" on the numeral identification task if they identified 6 or more numerals

### Index of Supplementary Materials

- Sarnecka, B. W., Negen, J., Scalise, N. R., Goldman, M. C., & Rouder, J. N. (2022). *Supplementary materials to "The real preschoolers of Orange County: Early number learning in a diverse group of children"* [Research data and stimuli]. OSF. <https://osf.io/tqrng/>
- Sarnecka, B. W., Negen, J., Scalise, N. R., Goldman, M. C., & Rouder, J. N. (2023). *Supplementary materials to "The real preschoolers of Orange County: Early number learning in a diverse group of children"* [Additional information]. PsychOpen GOLD. <https://doi.org/10.23668/psycharchives.12534>

## References

- Almoammer, A., Sullivan, J., Donlan, C., Marušič, F., Žaucer, R., O'Donnell, T., & Barner, D. (2013). Grammatical morphology as a source of early number word meanings. *Proceedings of the National Academy of Sciences*, *110*(46), 18448–18453. <https://doi.org/10.1073/pnas.1313652110>
- Anobile, G., Cicchini, G. M., & Burr, D. C. (2012). Linear mapping of numbers onto space requires attention. *Cognition*, *122*(3), 454–459. <https://doi.org/10.1016/j.cognition.2011.11.006>
- Baroody, A. J., & Price, J. (1983). The development of the number-word sequence in the counting of three-year-olds. *Journal for Research in Mathematics Education*, *14*(5), 361–368. <https://doi.org/10.2307/748681>
- Bialystok, E. (2009). Bilingualism: The good, the bad, and the indifferent. *Bilingualism: Language and Cognition*, *12*(1), 3–11. <https://doi.org/10.1017/S1366728908003477>
- Bialystok, E. (2011). Reshaping the mind: The benefits of bilingualism. *Canadian Journal of Experimental Psychology / Revue Canadienne de Psychologie Expérimentale*, *65*(4), 229–235. <https://doi.org/10.1037/a0025406>
- Bialystok, E., & Codd, J. (1997). Cardinal limits: Evidence from language awareness and bilingualism for developing concepts of number. *Cognitive Development*, *12*(1), 85–106. [https://doi.org/10.1016/S0885-2014\(97\)90031-9](https://doi.org/10.1016/S0885-2014(97)90031-9)
- Bialystok, E., Luk, G., Peets, K. F., & Yang, S. (2010). Receptive vocabulary differences in monolingual and bilingual children. *Bilingualism: Language and Cognition*, *13*(4), 525–531. <https://doi.org/10.1017/S1366728909990423>
- Bialystok, E., & Martin, M. M. (2004). Attention and inhibition in bilingual children: Evidence from the dimensional change card sort task. *Developmental Science*, *7*(3), 325–339. <https://doi.org/10.1111/j.1467-7687.2004.00351.x>
- Brannon, E. M., & Merritt, D. J. (2011). Chapter 14—Evolutionary foundations of the approximate number system. In S. Dehaene & E. M. Brannon (Eds.), *Space, time and number in the brain* (pp. 207–224). Academic Press. <https://doi.org/10.1016/B978-0-12-385948-8.00014-1>
- Briars, D., & Siegler, R. S. (1984). A featural analysis of preschoolers' counting knowledge. *Developmental Psychology*, *20*(4), 607–618. <https://doi.org/10.1037/0012-1649.20.4.607>
- Bull, R., & Lee, K. (2014). Executive functioning and mathematics achievement. *Child Development Perspectives*, *8*(1), 36–41. <https://doi.org/10.1111/cdep.12059>
- Cankaya, O., LeFevre, J.-A., & Dunbar, K. (2014). The role of number naming systems and numeracy experiences in children's rote counting: Evidence from Turkish and Canadian children. *Learning and Individual Differences*, *32*, 238–245. <https://doi.org/10.1016/j.lindif.2014.03.016>
- Carey, S. (2009). *The origin of concepts*. Oxford University Press.

- Chard, D. J., Clarke, B., Baker, S., Otterstedt, J., Braun, D., & Katz, R. (2005). Using measures of number sense to screen for difficulties in mathematics: Preliminary findings. *Assessment for Effective Intervention, 30*(2), 3–14. <https://doi.org/10.1177/073724770503000202>
- Chen, Q., & Li, J. (2014). Association between individual differences in non-symbolic number acuity and math performance: A meta-analysis. *Acta Psychologica, 148*, 163–172. <https://doi.org/10.1016/j.actpsy.2014.01.016>
- Cheung, P., Rubenson, M., & Barner, D. (2017). To infinity and beyond: Children generalize the successor function to all possible numbers years after learning to count. *Cognitive Psychology, 92*, 22–36. <https://doi.org/10.1016/j.cogpsych.2016.11.002>
- Clarke, B., & Shinn, M. R. (2004). A preliminary investigation into the identification and development of early mathematics curriculum-based measurement. *School Psychology Review, 33*(2), 234–248. <https://doi.org/10.1080/02796015.2004.12086245>
- Coll, C. G. (2015). Editorial: Continuity and change in child development. *Child Development, 86*(1), 7–9. <https://doi.org/10.1111/cdev.12353>
- Data USA. (2022, April). *Cities and places*. <https://datausa.io/>
- Daubert, E. N., & Ramani, G. B. (2019). Math and memory in bilingual preschoolers: The relations between bilingualism, working memory, and numerical knowledge. *Journal of Cognition and Development, 20*(3), 314–333. <https://doi.org/10.1080/15248372.2019.1565536>
- Davidson, K., Eng, K., & Barner, D. (2012). Does learning to count involve a semantic induction? *Cognition, 123*(1), 162–173. <https://doi.org/10.1016/j.cognition.2011.12.013>
- Dehaene, S. (2001). Précis of the number sense. *Mind & Language, 16*(1), 16–36. <https://doi.org/10.1111/1468-0017.00154>
- Dehaene, S. (2011). *The number sense: How the mind creates mathematics* (Revised and updated ed.). Oxford University Press.
- Duncan, G. J., Dowsett, C. J., Claessens, A., Magnuson, K., Huston, A. C., Klebanov, P., Pagani, L. S., Feinstein, L., Engel, M., Brooks-Gunn, J., Sexton, H., Duckworth, K., & Japel, C. (2007). School readiness and later achievement. *Developmental Psychology, 43*(6), 1428–1446. <https://doi.org/10.1037/0012-1649.43.6.1428>
- Dunn, L. M., & Dunn, L. M. (1997). *Peabody picture vocabulary test-III*. American Guidance Service.
- Dunn, L. M., Lugo, D. E., Padilla, E. R., & Dunn, L. M. (1986). *Test de Vocabulario en Imagenes Peabody (TVIP)*. Pearson.
- Emmorey, K., Luk, G., Pyers, J. E., & Bialystok, E. (2008). The source of enhanced cognitive control in bilinguals: Evidence from bimodal bilinguals. *Psychological Science, 19*(12), 1201–1206. <https://doi.org/10.1111/j.1467-9280.2008.02224.x>
- Fazio, L. K., Bailey, D. H., Thompson, C. A., & Siegler, R. S. (2014). Relations of different types of numerical magnitude representations to each other and to mathematics achievement. *Journal of Experimental Child Psychology, 123*, 53–72. <https://doi.org/10.1016/j.jecp.2014.01.013>
- Feigenson, L., Dehaene, S., & Spelke, E. (2004). Core systems of number. *Trends in Cognitive Sciences, 8*(7), 307–314. <https://doi.org/10.1016/j.tics.2004.05.002>
- Frye, D., Braisby, N., Lowe, J., Maroudas, C., & Nicholls, J. (1989). Young children's understanding of counting and cardinality. *Child Development, 60*(5), 1158–1171. <https://doi.org/10.2307/1130790>
- Fuson, K. C. (1988). *Children's counting and concepts of number* (Vol. 14). Springer.
- Fuson, K. C. (1992). Relationships between counting and cardinality from age 2 to age 8. In *Pathways to number: Children's developing numerical abilities* (pp. 127–149). Lawrence Erlbaum Associates.
- Geary, D. C., & Hoard, M. K. (2005). Learning disabilities in arithmetic and mathematics: Theoretical and empirical perspectives. In J. I. D. Campbell (Ed.), *Handbook of mathematical cognition* (pp. 253–267). Psychology Press.
- Geary, D. C., Hoard, M. K., Nugent, L., & Byrd-Craven, J. (2008). Development of number line representations in children with mathematical learning disability. *Developmental Neuropsychology, 33*(3), 277–299. <https://doi.org/10.1080/87565640801982361>
- Geary, D. C., & vanMarle, K. (2016). Young children's core symbolic and nonsymbolic quantitative knowledge in the prediction of later mathematics achievement. *Developmental Psychology, 52*(12), 2130–2144. <https://doi.org/10.1037/dev0000214>
- Geary, D. C., vanMarle, K., Chu, F. W., Rouder, J., Hoard, M. K., & Nugent, L. (2018). Early conceptual understanding of cardinality predicts superior school-entry number-system knowledge. *Psychological Science, 29*(2), 191–205. <https://doi.org/10.1177/0956797617729817>
- Gennetian, L. A., Rodrigues, C., Hill, H. D., & Morris, P. (2015). *Income instability in the lives of Hispanic children* (Publication #2015-47). National Research Center on Hispanic Children & Families. <http://www.childtrends.org/wp-content/uploads/2015/12/2015-47Hisp-Ctr-Income-Instability.pdf>

- Goldman, M. C., Negen, J., & Sarnecka, B. W. (2014). Are bilingual children better at ignoring perceptually misleading information? A novel test. *Developmental Science*, *17*(6), 956–964. <https://doi.org/10.1111/desc.12175>
- Griffin, S. A., Case, R., & Siegler, R. S. (1994). Rightstart: Providing the central conceptual prerequisites for first formal learning of arithmetic to students at risk for school failure. In K. McGilly (Ed.), *Classroom lessons: Integrating cognitive theory and classroom practice* (pp. 25–49). The MIT Press.
- Gunderson, E. A., Ramirez, G., Beilock, S. L., & Levine, S. C. (2012). The relation between spatial skill and early number knowledge: The role of the linear number line. *Developmental Psychology*, *48*(5), 1229–1241. <https://doi.org/10.1037/a0027433>
- Henrich, J., Heine, S. J., & Norenzayan, A. (2010). Most people are not WEIRD. *Nature*, *466*, Article 29. <https://doi.org/10.1038/466029a>
- Honoré, N., & Noël, M.-P. (2016). Improving preschoolers' arithmetic through number magnitude training: The impact of non-symbolic and symbolic training. *PLOS ONE*, *11*(11), Article e0166685. <https://doi.org/10.1371/journal.pone.0166685>
- Izard, V., Sann, C., Spelke, E. S., & Streri, A. (2009). Newborn infants perceive abstract numbers. *Proceedings of the National Academy of Sciences*, *106*(25), 10382–10385. <https://doi.org/10.1073/pnas.0812142106>
- Jordan, N. C., Kaplan, D., Locuniak, M. N., & Ramineni, C. (2007). Predicting first-grade math achievement from developmental number sense trajectories. *Learning Disabilities Research & Practice*, *22*(1), 36–46. <https://doi.org/10.1111/j.1540-5826.2007.00229.x>
- Jordan, N. C., Kaplan, D., Nabors Oláh, L., & Locuniak, M. N. (2006). Number sense growth in kindergarten: A longitudinal investigation of children at risk for mathematics difficulties. *Child Development*, *77*(1), 153–175. <https://doi.org/10.1111/j.1467-8624.2006.00862.x>
- Jordan, N. C., Kaplan, D., Ramineni, C., & Locuniak, M. N. (2009). Early math matters: Kindergarten number competence and later mathematics outcomes. *Developmental Psychology*, *45*(3), 850–867. <https://doi.org/10.1037/a0014939>
- Jordan, N. C., & Levine, S. C. (2009). Socioeconomic variation, number competence, and mathematics learning difficulties in young children. *Developmental Disabilities Research Reviews*, *15*(1), 60–68. <https://doi.org/10.1002/ddrr.46>
- Knudsen, B., Fischer, M. H., Henning, A., & Aschersleben, G. (2015). The development of Arabic digit knowledge in 4- to 7-year-old children. *Journal of Numerical Cognition*, *1*(1), 21–37. <https://doi.org/10.5964/jnc.v1i1.4>
- Kolkman, M. E., Kroesbergen, E. H., & Leseman, P. P. M. (2013). Early numerical development and the role of non-symbolic and symbolic skills. *Learning and Instruction*, *25*, 95–103. <https://doi.org/10.1016/j.learninstruc.2012.12.001>
- LeFevre, J.-A., Fast, L., Skwarchuk, S.-L., Smith-Chant, B. L., Bisanz, J., Kamawar, D., & Penner-Wilger, M. (2010). Pathways to mathematics: Longitudinal predictors of performance. *Child Development*, *81*(6), 1753–1767. <https://doi.org/10.1111/j.1467-8624.2010.01508.x>
- Leibovich, T., & Ansari, D. (2016). The symbol-grounding problem in numerical cognition: A review of theory, evidence, and outstanding questions. *Canadian Journal of Experimental Psychology/Revue Canadienne de Psychologie Expérimentale*, *70*(1), 12–23. <https://doi.org/10.1037/cep0000070>
- Li, P., Le Corre, M., Shui, R., Jia, G., & Carey, S. (2003, October). *Effects of plural syntax on number word learning: A cross-linguistic study*. Boston University Conference on Language Development (BUCLD), Boston, MA, USA.
- Libertus, M. E., & Brannon, E. M. (2009). Behavioral and neural basis of number sense in infancy. *Current Directions in Psychological Science*, *18*(6), 346–351. <https://doi.org/10.1111/j.1467-8721.2009.01665.x>
- Maertens, B., De Smedt, B., Sasanguie, D., Elen, J., & Reynvoet, B. (2016). Enhancing arithmetic in pre-schoolers with comparison or number line estimation training: Does it matter? *Learning and Instruction*, *46*, 1–11. <https://doi.org/10.1016/j.learninstruc.2016.08.004>
- Marchand, E., Wade, S., Sullivan, J., & Barner, D. (2020). Language-specific numerical estimation in bilingual children. *Journal of Experimental Child Psychology*, *197*, Article 104860. <https://doi.org/10.1016/j.jecp.2020.104860>
- Merkley, R., & Ansari, D. (2016). Why numerical symbols count in the development of mathematical skills: Evidence from brain and behavior. *Current Opinion in Behavioral Sciences*, *10*, 14–20. <https://doi.org/10.1016/j.cobeha.2016.04.006>
- Miller, E. B. (2016). How Spanish is used in head start: Observational evidence from four classrooms. *Head Start Dialog*, *19*(3), 1–26. <https://journals.uncc.edu/dialog/article/view/539>
- Moore, A. M., vanMarle, K., & Geary, D. C. (2016). Kindergartners' fluent processing of symbolic numerical magnitude is predicted by their cardinal knowledge and implicit understanding of arithmetic 2 years earlier. *Journal of Experimental Child Psychology*, *150*, 31–47. <https://doi.org/10.1016/j.jecp.2016.05.003>
- Morey, R. D., & Rouder, J. N. (2011). Bayes factor approaches for testing interval null hypotheses. *Psychological Methods*, *16*(4), 406–419. <https://doi.org/10.1037/a0024377>

- Morey, R. D., & Rouder, J. N. (2018). *BayesFactor: Computation of Bayes Factors for Common Designs* (R package version 0.9.12–4.2) [Computer software]. <https://cran.r-project.org/package=BayesFactor>
- Mussolin, C., Nys, J., Leybaert, J., & Content, A. (2012). Relationships between approximate number system acuity and early symbolic number abilities. *Trends in Neuroscience and Education, 1*(1), 21–31. <https://doi.org/10.1016/j.tine.2012.09.003>
- National Research Council. (2009). *Mathematics learning in early childhood: Paths toward excellence and equity*. The National Academies Press. [http://www.nap.edu/catalog.php?record\\_id=12519](http://www.nap.edu/catalog.php?record_id=12519)
- Negen, J., & Sarnecka, B. W. (2011). Analogue magnitudes and knower-levels: Re-visiting the variability argument. In S. Ohlsson & R. Catrambone (Eds.), *Proceedings of the 32nd Annual Meeting of the Cognitive Science Society* (pp. 1252–1257). Cognitive Science Society, Red Hook, NY, USA. <https://escholarship.org/uc/item/3gn0x904>
- Negen, J., & Sarnecka, B. W. (2012). Number-concept acquisition and general vocabulary development. *Child Development, 83*(6), 2019–2027. <https://doi.org/10.1111/j.1467-8624.2012.01815.x>
- Negen, J., & Sarnecka, B. W. (2015). Is there really a link between exact-number knowledge and approximate number system acuity in young children? *British Journal of Developmental Psychology, 33*(1), 92–105. <https://doi.org/10.1111/bjdp.12071>
- Negen, J., Sarnecka, B. W., & Lee, M. D. (2012). An Excel sheet for inferring children's number-knower levels from give-N data. *Behavior Research Methods, 44*(1), 57–66. <https://doi.org/10.3758/s13428-011-0134-4>
- Nielsen, M., Haun, D., Kärtner, J., & Legare, C. H. (2017). The persistent sampling bias in developmental psychology: A call to action. *Journal of Experimental Child Psychology, 162*, 31–38. <https://doi.org/10.1016/j.jecp.2017.04.017>
- Piantadosi, S. T., Jara-Ettinger, J., & Gibson, E. (2014). Children's learning of number words in an indigenous farming-foraging group. *Developmental Science, 17*(4), 553–563. <https://doi.org/10.1111/desc.12078>
- Piazza, M., Facoetti, A., Trussardi, A. N., Berteletti, I., Conte, S., Lucangeli, D., Dehaene, S., & Zorzi, M. (2010). Developmental trajectory of number acuity reveals a severe impairment in developmental dyscalculia. *Cognition, 116*(1), 33–41. <https://doi.org/10.1016/j.cognition.2010.03.012>
- Poulin-Dubois, D., Blaye, A., Coutya, J., & Bialystok, E. (2011). The effects of bilingualism on toddlers' executive functioning. *Journal of Experimental Child Psychology, 108*(3), 567–579. <https://doi.org/10.1016/j.jecp.2010.10.009>
- Purpura, D. J., Baroody, A. J., & Lonigan, C. J. (2013). The transition from informal to formal mathematical knowledge: Mediation by numeral knowledge. *Journal of Educational Psychology, 105*(2), 453–464. <https://doi.org/10.1037/a0031753>
- Purpura, D. J., & Napoli, A. R. (2015). Early numeracy and literacy: Untangling the relation between specific components. *Mathematical Thinking and Learning, 17*(2–3), 197–218. <https://doi.org/10.1080/10986065.2015.1016817>
- Ramani, G. B., & Siegler, R. S. (2008). Promoting broad and stable improvements in low-income children's numerical knowledge through playing number board games. *Child Development, 79*(2), 375–394. <https://doi.org/10.1111/j.1467-8624.2007.01131.x>
- Reynvoet, B., & Sasanguie, D. (2016). The Symbol Grounding Problem revisited: A thorough evaluation of the ANS mapping account and the proposal of an alternative account based on symbol–symbol associations. *Frontiers in Psychology, 7*, Article 1581. <https://doi.org/10.3389/fpsyg.2016.01581>
- Rouder, J. N., & Morey, R. D. (2012). Default Bayes factors for model selection in regression. *Multivariate Behavioral Research, 47*(6), 877–903. <https://doi.org/10.1080/00273171.2012.734737>
- Rouder, J. N., Speckman, P. L., Sun, D., Morey, R. D., & Iverson, G. (2009). Bayesian *t* tests for accepting and rejecting the null hypothesis. *Psychonomic Bulletin & Review, 16*(2), 225–237. <https://doi.org/10.3758/PBR.16.2.225>
- RStudio Team. (2020). *Stats* (R package version 1.2.1335) [Computer software]. RStudio PBC. <http://www.rstudio.com>
- Sarnecka, B. W. (2016). How numbers are like the earth (and unlike faces, loitering or knitting). In D. Barner & A. S. Baron (Eds.), *Core knowledge and conceptual change* (pp. 151–170). Oxford University Press.
- Sarnecka, B. W., & Carey, S. (2008). How counting represents number: What children must learn and when they learn it. *Cognition, 108*(3), 662–674. <https://doi.org/10.1016/j.cognition.2008.05.007>
- Sarnecka, B. W., Kamenskaya, V. G., Yamana, Y., Ogura, T., & Yudovina, Y. B. (2007). From grammatical number to exact numbers: Early meanings of 'one', 'two', and 'three' in English, Russian, and Japanese. *Cognitive Psychology, 55*(2), 136–168. <https://doi.org/10.1016/j.cogpsych.2006.09.001>
- Sarnecka, B. W., & Lee, M. D. (2009). Levels of number knowledge during early childhood. *Journal of Experimental Child Psychology, 103*(3), 325–337. <https://doi.org/10.1016/j.jecp.2009.02.007>
- Sarnecka, B. W., & Wright, C. E. (2013). The idea of an exact number: Children's understanding of cardinality and equinumerosity. *Cognitive Science, 37*(8), 1493–1506. <https://doi.org/10.1111/cogs.12043>

- Sasanguie, D., Smedt, B. D., Defever, E., & Reynvoet, B. (2012). Association between basic numerical abilities and mathematics achievement. *British Journal of Developmental Psychology*, *30*(2), 344–357. <https://doi.org/10.1111/j.2044-835X.2011.02048.x>
- Schneider, M., Beeres, K., Coban, L., Merz, S., Susan Schmidt, S., Stricker, J., & De Smedt, B. (2017). Associations of non-symbolic and symbolic numerical magnitude processing with mathematical competence: A meta-analysis. *Developmental Science*, *20*(3), Article e12372. <https://doi.org/10.1111/desc.12372>
- Schneider, M., Merz, S., Stricker, J., De Smedt, B., Torbeyns, J., Verschaffel, L., & Luwel, K. (2018). Associations of number line estimation with mathematical competence: A meta-analysis. *Child Development*, *89*(5), 1467–1484. <https://doi.org/10.1111/cdev.13068>
- Schneider, R. M., Sullivan, J., Guo, K., & Barner, D. (2020). What counts? Sources of knowledge in children's acquisition of the successor function. *Child Development*, *92*(4), e476–e492. <https://doi.org/10.1111/cdev.13524>
- Siegler, R. S. (2016). Magnitude knowledge: The common core of numerical development. *Developmental Science*, *19*(3), 341–361. <https://doi.org/10.1111/desc.12395>
- Siegler, R. S., & Ramani, G. B. (2009). Playing linear number board games—But not circular ones—Improves low-income preschoolers' numerical understanding. *Journal of Educational Psychology*, *101*(3), 545–560. <https://doi.org/10.1037/a0014239>
- Slaughter, V., Itakura, S., Kutsuki, A., & Siegal, M. (2011). Learning to count begins in infancy: Evidence from 18 month olds' visual preferences. *Proceedings of the Royal Society B: Biological Sciences*, *278*(1720), 2979–2984. <https://doi.org/10.1098/rspb.2010.2602>
- Slusser, E. B., & Sarnecka, B. W. (2011). Find the picture of eight turtles: A link between children's counting and their knowledge of number word semantics. *Journal of Experimental Child Psychology*, *110*(1), 38–51. <https://doi.org/10.1016/j.jecp.2011.03.006>
- Spaepen, E., Gunderson, E. A., Gibson, D., Goldin-Meadow, S., & Levine, S. C. (2018). Meaning before order: Cardinal principle knowledge predicts improvement in understanding the successor principle and exact ordering. *Cognition*, *180*, 59–81. <https://doi.org/10.1016/j.cognition.2018.06.012>
- Starkey, P., Klein, A., & Wakeley, A. (2004). Enhancing young children's mathematical knowledge through a pre-kindergarten mathematics intervention. *Early Childhood Research Quarterly*, *19*(1), 99–120. <https://doi.org/10.1016/j.ecresq.2004.01.002>
- Starr, A., Libertus, M. E., & Brannon, E. M. (2013). Number sense in infancy predicts mathematical abilities in childhood. *Proceedings of the National Academy of Sciences*, *110*(45), 18116–18120. <https://doi.org/10.1073/pnas.1302751110>
- U.S. Census Bureau. (2019). *American Communities Survey*. <https://data.census.gov>
- van Marle, K., Chu, F. W., Li, Y., & Geary, D. C. (2014). Acuity of the approximate number system and preschoolers' quantitative development. *Developmental Science*, *17*(4), 492–505. <https://doi.org/10.1111/desc.12143>
- Walters, D. (2015, November 5). Orange County is home to 108 languages, census finds. *Orange County Register*. <https://www.ocregister.com/2015/11/05/orange-county-is-home-to-108-languages-census-finds>
- Wang, J., Halberda, J., & Feigenson, L. (2021). Emergence of the link between the approximate number system and symbolic math ability. *Child Development*, *92*(2), e186–e200. <https://doi.org/10.1111/cdev.13454>
- Watts, T. W., Duncan, G. J., Siegler, R. S., & Davis-Kean, P. E. (2014). What's past is prologue: Relations between early mathematics knowledge and high school achievement. *Educational Researcher*, *43*(7), 352–360. <https://doi.org/10.3102/0013189X14553660>
- Wynn, K. (1990). Children's understanding of counting. *Cognition*, *36*(2), 155–193. [https://doi.org/10.1016/0010-0277\(90\)90003-3](https://doi.org/10.1016/0010-0277(90)90003-3)
- Wynn, K. (1992). Children's acquisition of the number words and the counting system. *Cognitive Psychology*, *24*(2), 220–251. [https://doi.org/10.1016/0010-0285\(92\)90008-P](https://doi.org/10.1016/0010-0285(92)90008-P)
- Xu, F., Spelke, E. S., & Goddard, S. (2005). Number sense in human infants. *Developmental Science*, *8*(1), 88–101. <https://doi.org/10.1111/j.1467-7687.2005.00395.x>



*Journal of Numerical Cognition* (JNC) is an official journal of the Mathematical Cognition and Learning Society (MCLS).



leibniz-psychology.org

PsychOpen GOLD is a publishing service by Leibniz Institute for Psychology (ZPID), Germany.